Multivariate data analysis of near-infrared spectra of cultivation medium powders for mammalian cells

É. Szabó, S. Gergely and A. Salgó*

Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics. E-mail: salgo@mail.bme.hu

Nowadays, qualification and control of medium formulations is performed based on simple methods (e.g. pH and osmolality measurement of media solutions), expensive and time-consuming cell culture tests, and quantification of some critical compounds by liquid chromatography. Besides the traditional medium qualification tools, relatively new spectroscopic techniques such as fluorescence spectroscopy, nuclear magnetic resonance, Raman and NIR spectroscopies or a combination of these techniques are increasingly being applied for cultivation medium powder investigation. A chemically defined cultivation medium powder for Chinese hamster ovary (CHO) cell cultivation was investigated in this study, regarding its response to heat treatments with different temperatures (30 °C, 50 °C and 70 °C). The heat treatments were performed according to a design of experiments (DoE) approach. Spectra of the control and the treated powders were collected to compare the sample groups using a dispersive near-infrared (NIR) and a Fourier-transform near-infrared (FT-NIR) spectrometer. Multivariate data analysis including unsupervised (cluster analysis, principal component analysis, polar qualification system) and supervised classification methods (linear discriminant analysis, soft independent modelling of class analogies and partial least squares discriminant analysis) were employed to investigate the monitoring capability of near-infrared spectroscopy for qualification of cultivation medium powder stored at different temperatures, to identify the treatment-induced variations in the samples, and compare the efficiency of spectrometers with distinct optical arrangements (i.e. dispersive and Fourier transformed spectrometers). During heat treatment, cultivation medium powders went through spectroscopically recognisable changes. Both NIR and FT-NIR analysis could separate samples according to the temperature set-points, irrespective of spectrometer attributes. In classification of samples cluster analysis and linear discriminant analysis shown the best results for both NIR and FT-NIR spectra.

Introductions

One of the most important groups of raw materials in biopharmaceutical production is cell culture medium powder. The chemical composition of these cultivation medium powders is a critical factor for cell culture process performance,[1] as it directly influences cell proliferation and productivity and indirectly product quality. Cultivation medium powders for mammalian cells contain even 60–80 chemical components, these are usually inorganic salts, carbohydrates, amino acids, vitamins, peptones or proteins and other miscellaneous compounds.[2] Additionally, media powder production involves multiple process steps, which can contribute to the quality variability of formulations.

Nowadays qualification and control of medium powders are performed based on time-consuming, labour intensive and expensive cell culture tests together with common physico–chemical measurements of dissolved medium powders (e.g. pH and osmolality). Qualification of critical compounds also can be analysed by liquid chromatography. Besides the traditional cultivation medium qualification tools, spectroscopic techniques such as fluorescence, Raman and near-infrared spectroscopies, are increasingly being applied for cultivation medium powder investigation.[3] NIR spectroscopy is a fast and non-destructive analytical tool, and combined with multivariate data analysis, NIR spectra provide physical infor-

Correspondence
A. Salgó (salgo@mail.bme.hu)
doi: 10.1255/nir2017.143
Citation: É. Szabó, S. Gergely and A. Salgó, "Multivariate data analysis of near-infrared spectra of cultivation medium powders for mammalian cells", in *Proc. 18th Int. Conf. Near Infrared Spectrosc.*, Ed by S.B. Engelsen, K.M. Sørensen and F. van den Berg. IM Publications Open, Chichester, pp. 143–150 (2019). https://doi.org/10.1255/nir2017.143

ISBN: 978-1-906715-27-4

mation and chemical fingerprints of the raw materials, thus give rise an analytical tool to distinguish high-quality and damaged media powders.[4]

Cultivation medium powders are stable and normally kept at 2–8 °C and detecting the effect of harmful higher temperatures is crucial, since shorter heat shocks can occur during improper powder handling (e.g. transport or storage). Hakemeyer *et al.*[5] applied NIR and two-dimensional fluorescence spectroscopy for monitoring cultivation medium powder formulations after long-term (7–84 days) room temperature (22 ± 4 °C) storage. Previous publications of the authors introduced the monitoring capability of NIR and Fourier transform NIR (FT-NIR) spectroscopy for a cultivation medium powder stored at 30 °C, 50 °C and 70 °C. NIR spectroscopy could separate samples according to temperature set-points of the heat treatments successfully, indicating that NIR spectroscopy is a promising tool for cultivation medium powder qualification, which can be used to control cell culture raw material variability in a fast and cost effective way.[6] These studies prove that NIR spectroscopy could be used to discriminate the control and heat-treated samples and investigation of the effects of higher temperatures or other treatments can be feasible using experimentally treated samples.

Several algorithms for classification tasks have been described in the literature, which can be divided into two categories: the unsupervised (e.g. cluster analysis, principal component analysis, polar qualification system) and the supervised methods (like linear discriminant analysis, soft independent modelling of class analogies and partial least squares discriminant analysis). In unsupervised pattern recognition, algorithms do not use *a priori* information about class membership of samples. The modelling is based on the input variables (i.e. x-variables), therefore, only spectral data are required, and the samples can be categorised based on the pattern obtained in the results. In contrast, supervised classification algorithms use the labels of class membership of samples to make models, thus these methods require both spectral (x-variables) and reference data (y-variables).[7,8]

Cluster analysis (CA) or hierarchical cluster analysis (HCA) is an unsupervised multivariate technique, which creates a tree-shaped data chart of samples, called a dendrogram. This dendrogram organises data into different sub-categories, which are then sub-divided to lower levels allow the examination of grouping relations between samples. The frequently used algorithms of CA use agglomerative techniques, which define distances or similarity between samples (e.g. Euclidean distance, correlation), then form clusters using amalgamation rules (e.g. single linkage, complete linkage, Ward's method). Different distance measures combined with different amalgamation rules give different results, therefore, these techniques must be selected according to the type of data set.[9] Among unsupervised methods, principal component analysis (PCA) is one of the most widespread. PCA can be used to reduce the number of variables, since this technique calculates orthogonal latent variables (i.e. principal components) from the original data matrix, which can describe the variability of samples to the highest possible degree. According to the results, the multivariate data set can be projected into a low dimensional space for visualisation, and the weights of original variables also can be interpreted like a spectrum.[10] The polar qualification system (PQS) method is different from the ones mentioned above. PQS reduces the large-scale spectral data based on geometrical considerations. The spectra are converted from a Cartesian coordinate system into a polar coordinate system, and they can be represented by the centre of the resulting shapes (i.e. quality points) in the polar coordinate system. The advantage of this method is its simplicity and easy representability.[11]

Linear discriminant analysis (LDA) is one of the most widely used supervised classification methods. LDA focuses on discrimination, since select those directions from the high dimensional space, which achieve maximum separation among different classes. LDA is a linear and parametric technique, and uses Euclidean distance measurement to classify unknown samples. The disadvantage of this method that it requires the normal distribution of statistical parameters for the correct classification of samples and often results in over-fitted model that can perfectly fit the training data while performing poorly on unknown samples.[12] Soft independent modelling of class analogies (SIMCA) is a supervised classification technique that considers each class separately. SIMCA performs a separate PCA for each class as a model. The unknown samples are fitted into these models and classified as the member of class if they are in the confidence interval of the certain PCA model. This technique allows a new sample to be classified into more classes or none of them.[13] Partial least squares discriminant analysis (PLS-DA) is also a linear and parametric method, which implements PLS analysis of a spectral matrix and the categorical reference vector. For binary data, the refer-

ence vector contains 0 or 1 values according to the class membership. Unknown samples result in response values between 0 and 1 in the prediction, and the threshold (usually 0.5) can be used to decide the class membership for these samples. Multiclass problems can be solved using several binary models or a single model including all groups with a separate binary variable (i.e. dummy variables). Likewise, LDA, this technique also can result in over-fitted models.[14]

The aims of this study are to investigate the monitoring capability of NIR spectroscopy for qualification of cultivation medium powders stored at different temperatures, to identify the treatment-induced variations in the samples, and compare the efficiency of spectrometers with distinct optical arrangements (i.e. dispersive and Fourier transformed spectrometers) using different multivariate data analysis techniques including unsupervised (CA, PCA, PQS) and supervised classification methods (LDA, SIMCA, PLS-DA).

Materials and methods

Samples and spectral measurements

Chemically defined PowerCHO-2 (Lonza, Walkersville, MD, USA) cultivation medium powder was treated in 30 °C, 50 °C and 70 °C. To increase the variability of treated samples, different exposure times (1 hour, 7 hours and 13 hours) were included as a random factor in the full factorial design of experiment. Control samples were stored at 4 °C. Cell cultivation experiments with Chinese hamster ovary cell line indicated that the heat-induced changes of the medium affect cell culture performance, as samples treated at higher temperature (≥50 °C) and duration (≥7 hours) resulted in VCDs below the predefined threshold.

NIR reflectance [log(1/R)] spectra of samples were scanned in diffuse reflectance mode using a Foss NIRSystems 6500 spectrometer (Foss NIRSystems, Inc., Silver Spring, MD, USA) equipped with a rapid content analyser (RCA). 32 scans with data interval of 2 nm were recorded for each spectrum in the wavelength rage of 1100–2498 nm.

FT-NIR reflectance spectra were recorded using a Spectrum 400 spectrometer (PerkinElmer, Inc., Waltham, MA, USA) with a NIR reflectance accessory (NIRA) which contains a gold-coated integrated sphere and an InGaAs detector. 32 scans with data interval of 2 cm^{-1} were recorded for each spectrum in the wavenumber range between 10,000 cm^{-1} and 4000 cm^{-1} (i.e. in the wavelength range between 1000 nm and 2500 nm).

Second derivatives have the capability to remove both baseline and linear trends in the spectra. The most common derivative techniques in spectral analysis are gap-segment and the Savitzky-Golay polynomial derivative methods. Both methods use a smoothing of the spectra prior to calculating the derivative in order to improve the signal-to-noise ratio in the corrected spectra. In the case of the gap-segment technique, a moving window with optimised segment and gap sizes is applied for smoothing.[15] In this study, second derivatives were calculated for all spectra with a gap of 1 point and a segment of 7 in the case of data processing of NIR spectra; correspondingly a gap of 1 point and a segment of 19 points for data processing of FT-NIR spectra using The Unscrambler X ver. 10.3 software (CAMO Software AS., Oslo, Norway).

PCA analysis of NIR and FT-NIR spectra showed that the combination tone region between 2200 nm and 2300 nm (4550 cm^{-1} and 3550 cm^{-1}) has the highest variability and accordingly this region were used for further investigations. Details of the sample preparation, cell cultivation experiments, NIR and FT-NIR measurements have been published previously.[6]

Multivariate data analyses

Second derivative NIR and FT-NIR spectra were analysed between 2200 nm and 2300 nm with data interval of 2 nm (i.e. 51 variables) and between 4550 cm^{-1} and 3550 cm^{-1} with a data interval of 2 cm^{-1} (i.e. 101 variables), respectively. Unsupervised methods were performed with all spectra. CA was calculated with squared Euclidean distance combined with Ward's method. PCA was accomplished based on the covariance matrix using the NIPALS algorithm. Two-dimensional score plots of samples were presented according to the first and second principal components (PC1 and PC2, respectively). Classes were identified with ellipses calculated from the two-dimensional normal distribution of the classes of spectra with confidence level 0.95. Both CA and PCA were executed in Statistica 64 ver. 13.2 software (Dell, Inc., Tulsa, OK, USA). PQS was performed using PQS32 ver. 1.37 software (Metrika R&D Co., Budapest, Hungary). Classes were also determined with ellipses of the classes of quality points of spectra on the two-dimensional polar system (quality plane).

For supervised classifications two sets of spectra were predefined: the test set including one spectrum from each sample and the training set including the remaining spectra. LDA models were calculated using a forward stepwise method and test samples were identified using three latent variables (i.e. roots). SIMCA analysis was performed based on the separate PCA models of classes. Classification of test samples were executed based on sample-to-model distances ($p = 0.05$). PLS-DA models with two latent variables (i.e. components) were built using the NIPALS algorithm with seven-fold cross-validation. Both LDA and PLS-DA were performed using Statistica 64 ver. 13.2 software (Dell, Inc., Tulsa, OK, USA), SIMCA was performed using The Unscrambler X ver. 10.3 software (CAMO Software AS., Oslo, Norway).

Results and discussions

Unsupervised analyses

To examine the treatment-induced variations in the samples, unsupervised multivariate methods were applied. The efficiency of CA, PCA and PQS can be compared on the basis of the patterns of classes. In CA the first four clusters comply with the classes of temperature of treatments (Figure 1). The class of 70 °C treated samples is the first cluster that diverges from the other samples with high linkage distance. Moreover, in the NIR dendrogram the cluster of 70 °C samples shows sub-groups according to the exposure times (1, 7 and 13 hours). Therefore, this class that had poor performance in cell cultivation experiments in the previous study,[6] can be reliably distinguished from the control samples and the samples treated in lower temperatures based on NIR or FT-NIR spectra. In case of NIR spectra, the second cluster is the 30 °C class, while FT-NIR spectra result in a second cluster with control samples. The latter is more favourable, since it makes it easier to distinguish control samples from the class of 30 °C and 50 °C samples. Although the clusters of 30 °C and 50 °C classes of the FT-NIR spectra has incorrectly classified some samples (two spectra from the 30 °C class locate in the class of 50 °C), in case of the dispersive NIR spectra all of these spectra locate in the right clusters.

While dendrograms of CA can be evaluated according to the distances of samples and clusters, PCA has the potential to present distances and orientations of samples in two- or three-dimensional space of PC scores. The two-dimensional score plots of PC1 and PC2 of NIR and FT-NIR spectra can be observed in Figure 2. On the plots, normal ellipses of data points of classes with confidence level 0.95 are presented to define areas for classes. Interestingly, the classes do not show linear, equidistance-type tendency according to the increased heat treatments. This can be explained by the non-linear nature of changes in the samples. While lower temperature (30 °C) treatment can cause significant physical conversion (e.g. particle size distribution), higher temperatures may damage the heat sensitive components of these media powders (e.g. vitamins).[6] In PCA, as in dendrograms of CA, the class of 70 °C spectra can be easily distinguished based on the score values of PC1. The control, 30 °C and 50 °C classes show overlap according to the score values of PC1 and PC2. In the case of NIR spectra, classes of control, 30 °C, and 50 °C samples show mutual overlap. In contrast, on the plot of FT-NIR just 30 °C and 50 °C classes show signifi-

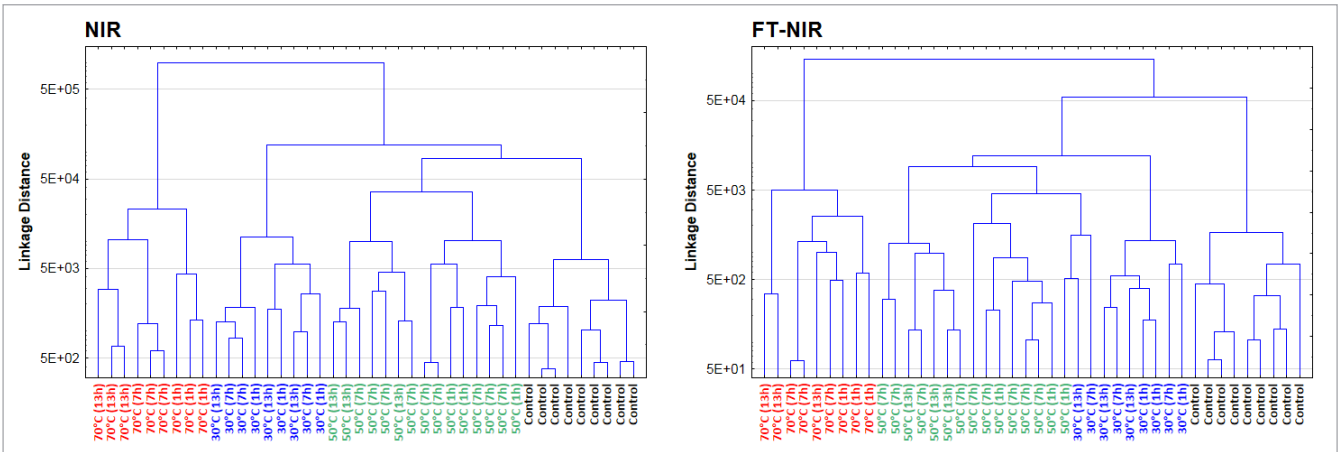

Figure 1. CA dendrograms of NIR and FT-NIR spectra of control and heat-treated cultivation medium powders.

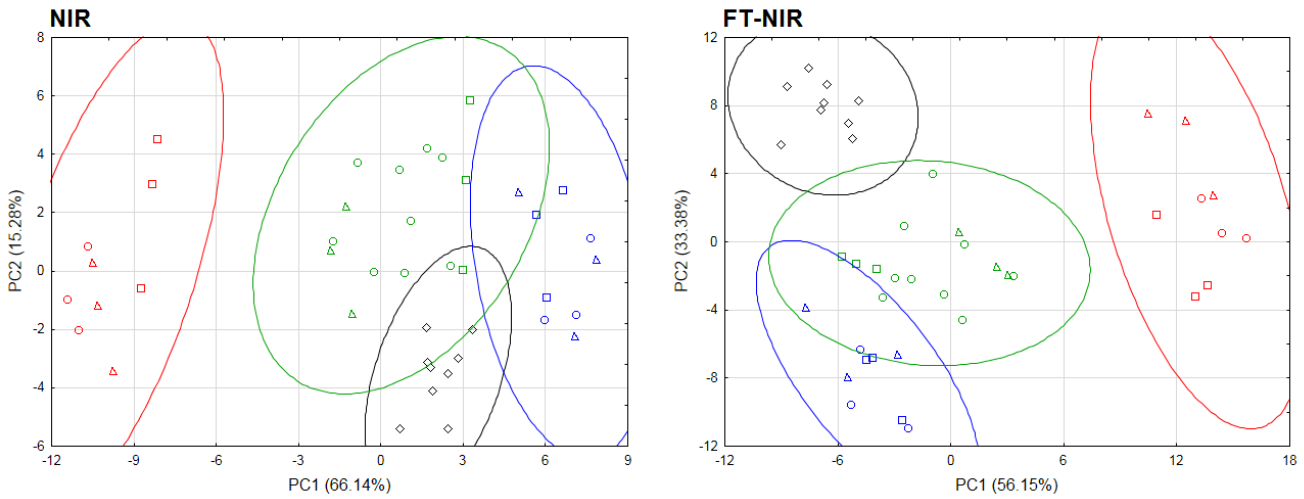

Figure 2. PCA score plots of NIR and FT-NIR spectra with normal ellipses of classes (confidence level: 0.95). (*Black diamond:* control samples, *blue:* 30 °C, *green:* 50 °C, *red:* 70 °C, *square:* 1 hour, *circle:* 7 hours, *triangle:* 13 hours exposure time.)

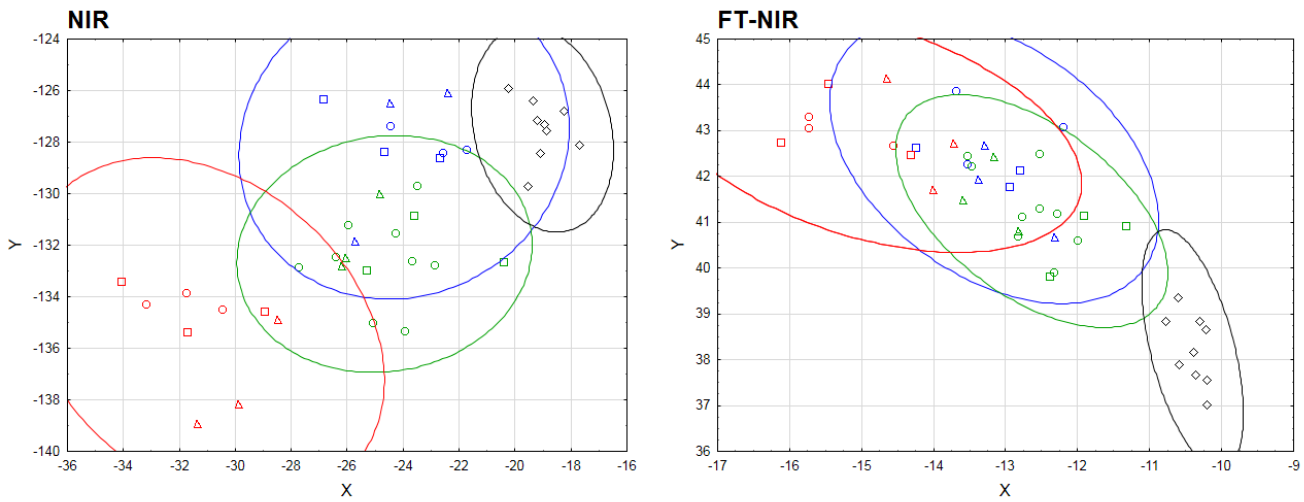

Figure 3. PQS quality plane of NIR and FT-NIR spectra with normal ellipses of classes (confidence level: 0.95). (*Black diamond:* control samples, *blue:* 30 °C, *green:* 50 °C, *red:* 70 °C, *square:* 1 hour, *circle:* 7 hours, *triangle:* 13 hours exposure time.)

cant overlap, and control class overlap only with the 50 °C class, and all samples of control class located to the right are of PC2 vs PC1 plane. These results agree with the dendrograms of CA: NIR spectra show smaller distance among control, 30 °C and 50 °C clusters, while FT-NIR spectra have a better separable control cluster, and more overlapped 30 °C and 50 °C clusters.

Results of PQS were also evaluated on the basis of two-dimensional normal ellipses of data points of classes on the PQS quality plane with confidence level 0.95 (Figure 3). In case of PQS, the classifications of samples are not appropriate. Dispersive NIR spectra result in overlapped clusters of each class and PQS analysis of FT-NIR spectra can separate only the control samples. The poor performance of this method can be ascribed to the simplicity of PQS, and the small differences among the spectra of classes, which cannot be detect based on geometrical transformations.

Supervised analyses

For supervised analyses, two sets of pre-treated spectra were predefined: the test set including one spectrum from each sample and the training set including the remaining spectra. In the training set, four classes were predefined: the control class and the three heat-treated classes (30 °C, 50 °C and 70 °C); and these classes were used to build LDA, SIMCA and PLS-DA models, which were compared on the basis of the numbers of correctly and incorrectly identified samples of the test set.

In LDA models three roots were computed. In the case of NIR spectra, all samples of the test set are correctly identified, but this method often results in over-fitted models, which cannot be precluded without independent samples. FT-NIR spectra also have good results; only one sample of the 30 °C class (1 hour) is incorrectly identified as a member of the 50 °C class.

SIMCA analysis has worse results than LDA: only test samples from the 50 °C class are perfectly classified on the basis of the distinct PCA models of classes. These results are surprising, since the above mentioned unsupervised PCA analyses have shown that the 50 °C class overlap with other classes. The better performance of SIMCA can be explained by the numbers of PCs: while only PC1 and PC2 were used to define the classes of samples, in SIMCA analyses the optimal number of PCs (up to five PCs) could be applied for the identification of test samples. Because SIMCA is a one-class modelling technique, samples may be classified into no classes or more than one class. There is no test sample which has been misclassified, all identified samples are in predicted to the right class. However, more samples cannot be classified into any classes of the NIR and FT-NIR models. In case of the NIR spectra a control sample, samples with 30 °C (with 1 and 13 hours treatment) and a sample from 70 °C class (with 13 hours treatment) cannot be identified as a member of any classes; these samples are categorised as outliers. SIMCA analysis of FT-NIR spectra also has outliers: 30 °C with 1 hour, 30 °C with 7 hours and 70 °C with 1-hour heat treatment. The first (30 °C, 1 hour) is identical with the sample, which has been incorrectly identified in the LDA of FT-NIR spectra.

PLS-DA analysis shows similar results in case of both NIR and FT-NIR spectra. Test samples from control, 30 °C and 70 °C classes can be identified properly, but all test samples from 50 °C class are misclassified. In case of the NIR spectra, these samples are identified as members of control, 30 °C or 70 °C class, while the PLS-DA model of FT-NIR spectra identified the 50 °C samples as members of 30 °C or 70 °C class. These results of FT-NIR are better than the results of PLS-DA of NIR spectra, since there is no incorrectly identified control samples.

Consequently, NIR and FT-NIR spectra have similar effectiveness during supervised and unsupervised analyses. A summary of the results can be obtained in Table 1. The best results are provided by CA and LDA in case of unsupervised and supervised multivariate methods, respectively. These techniques are able to identify the samples according to the temperature of heat treatment. In case of this data set, FT-NIR has better result in discriminating control samples from heat-treated samples.

Conclusions

Near-infrared spectra of heat-treated cultivation medium powders for CHO cell cultivation were investigated in this study, to compare the efficiency of different unsupervised and supervised classification methods and compare the monitoring capability of near-infrared spectrometers with dispersive and Fourier-transformed optical arrangements.

Table 1. Summary of results of applied multivariate methods based on NIR and FT-NIR spectra of control and heat-treated medium powders. (✓✓: no mistake in group; ✓: 1 or 2 mistakes; ×: 3 or more mistakes).

	Unsupervised methods						Supervised methods					
	CA		PCA		PQS		LDA		SIMCA		PLS-DA	
Group	NIR	FT-NIR	NIR	FT-NIR	NIR	FT-NIR	NIR	FT-NIR	NIR	FT-NIR	NIR	FT-NIR
Control	✓✓	✓✓	✓	✓✓	×	✓✓	✓✓	✓✓	✓	✓✓	✓✓	✓✓
30 °C	✓✓	✓	✓	×	×	×	✓✓	✓	✓	✓	✓✓	✓✓
50 °C	✓✓	✓✓	✓	✓	×	×	✓✓	✓✓	✓✓	✓✓	×	×
70 °C	✓✓	✓✓	✓✓	✓✓	✓	×	✓✓	✓✓	✓	✓	✓✓	✓✓

Samples were analysed using a dispersive and a Fourier-transformed spectrometer. After second derivative preprocessing of spectra, data points between 2200 nm and 2300 nm (4550 cm^{-1} and 3550 cm^{-1}) were used as variables for multivariate analyses including unsupervised (CA, PCA, PQS) and supervised classification methods (LDA, SIMCA, PLS-DA).

Media powders went through spectroscopically recognisable changes during heat treatment. Both NIR and FT-NIR analysis could separate control samples from heat-treated samples. In classification of samples, cluster analysis and linear discriminant analysis show the best results for both NIR and FT-NIR spectra. These techniques are able to distinguish samples according to the temperature set-points irrespective of spectrometer attributes. Consequently, differentiation of control and heat-treated samples may be achievable with different spectrometer attributes and various statistical methods.

The presented near-infrared spectroscopic method is a sensitive tool for medium powder qualification and will enable fast and easy assessment of raw material variability and could be a potential solution for rapid and high throughput qualification of raw materials in the pharmaceutical industry. These results should be completed with spectra of samples from other lots of this media powder and other media products to explore the lot-to-lot variability and the spectral differences among similar media powder products. The effects of improper powder handling during transport and storage can take the form of changing of humidity, particle size distribution or inhomogeneity of constituents. These kinds of anomalies can be easily monitored using near-infrared spectroscopy even with near-infrared chemical imaging.

References

1. A. Gilbert, Y. Huang and T. Ryll, "Identifying and eliminating cell culture process variability", *Pharm. Bioprocess.* **2(6),** 519–534 (2014). https://doi.org/10.4155/pbp.14.35
2. D. Jayme, T. Watanabe and T. Shimada, "Basal medium development for serum-free culture: a historical perspective", *Cytotechnology* **23(1-3),** 95–101 (1997). https://doi.org/10.1023/A:1007967602484
3. H.W. Lee, A. Christie, J. Xu and S. Yoon, "Data fusion-based assessment of raw materials in mammalian cell culture", *Biotechnol. Bioeng.* **109(11),** 2819–2828 (2012). https://doi.org/10.1002/bit.24548
4. A.O. Kirdar, G. Chen, J. Weidner and A.S. Rathore, "Application of near-infrared (NIR) spectroscopy for screening of raw materials used in the cell culture medium for the production of a recombinant therapeutic protein", *Biotechnol. Prog.* **26(2),** 527–531 (2010). https://doi.org/10.1002/btpr.329
5. C. Hakemeyer, U. Strauss, S. Werz, F. Folque and J.C. Menezes, "Near-infrared and two-dimensional fluorescence spectroscopy monitoring of monoclonal antibody fermentation media quality: aged media decreases cell growth", *Biotechnol. J.* **8(7),** 835–846 (2013). https://doi.org/10.1002/biot.201200355
6. É. Szabó, L. Párta, S. Gergely and A. Salgó, "Investigation of heat-treated cultivation medium for mammalian cells with near infrared spectroscopy", *J. Near Infrared Spectrosc.* **24(4),** 373–380 (2016). https://doi.org/10.1255/jnirs.1222
7. Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond and N. Jent, "A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies", *J. Pharm. Biomed. Anal.* **44(3),** 683–700 (2007). https://doi.org/10.1016/j.jpba.2007.03.023
8. N.L. Calvo, R.M. Maggio and T.S. Kaufman, "Characterization of pharmaceutically relevant materials at the solid state employing chemometrics methods", *J. Pharm. Biomed. Anal.* **147,** 538–564 (2018). https://doi.org/10.1016/j.jpba.2017.06.017
9. J.H. Ward Jr, "Hierarchical grouping to optimise an objective function", *J. Am. Stat. Assoc.* **58,** 236–244 (1963).
10. S. Wold, K. Esbensen and P. Geladi, "Principal component analysis", *Chemometr. Intell. Lab.* **2,** 37–52 (1987). https://doi.org/10.1016/0169-7439(87)80084-9
11. K.J. Kaffka and L.S. Gyarmati, "Investigating the Polar Qualification System", *J. Near Infrared Spectrosc.* **6,** A191–A200 (1998). https://doi.org/10.1255/jnirs.193
12. R.A. Fisher, "The use of multiple measurements in taxonomic problems", *Ann. Eugenics* **7,** 179–188 (1936).
13. S. Wold, "Pattern-recognition by means of disjoint principal components models", *Pattern Recogn.* **8,** 127–139 (1976).

14. M. Sjöström, S. Wold and B. Söderström, "PLS discriminant plots", in *Pattern Recognition in Practice II*, Ed by E.S. Gelsema and L.N. Kanal. Elsevier, Amsterdam, pp. 461–470 (1986).

15. Å. Rinnan, F. van den Berg and S.B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra", *Trends Anal. Chem.* **28(10),** 1201–1222 (2009). https://doi.org/10.1016/j.trac.2009.07.007